SECTION 3

MAKING USE SUPPORT DETERMINATIONS

This section presents EPA's recommended approaches to making use support decisions. Designated uses are assigned to individual waterbodies in a state's water quality standards. Types of designated uses include: aquatic life, fish consumption, recreational uses such as swimming, and drinking water. This guidance is drafted for wadeable streams and rivers. However, the approach is applicable to other types of waterbodies, as well.

3.1 ITFM Recommendations for Monitoring

The Intergovernmental Task Force on Monitoring Water Quality (ITFM) was formed in 1992 to develop recommendations on monitoring to achieve more comparable and scientifically defensible information, interpretations, and evaluations of water-quality conditions across the nation. The ITFM comprised both Federal and State agencies responsible for monitoring and assessment programs as well as an associated advisory committee including municipalities, academia, industry, etc. (ITFM 1995). The ITFM subsequently developed a model for stream monitoring for different types of designated uses based on a combination of biological, physical, and chemical monitoring (Figure 3-1). The model defines the relationship between parameters that directly measure the condition of the biotic community and its response over time to stressors, such as fish and benthic macroinvertebrate indices, and parameters that measure either stressors or exposure of organisms to stressors, such as levels of pH, nutrients, and toxicants. For streams, EPA recommends that States incorporate ITFM's suite of parameters in their monitoring programs for evaluating attainment of designated uses. These are general recommendations to consider when developing and revising monitoring programs. For example, monitoring for aquatic life use would include the base monitoring program parameters in the box--community level biological data from at least two assemblages, habitat, and physical/chemical field parameters—plus ionic strength, nutrients, and toxicants in water and sediment.

The ITFM in May 1997 became a permanent National Water Quality Monitoring Council to facilitate, among other tasks, the development and implementation of the recommendations on specific methods for measuring

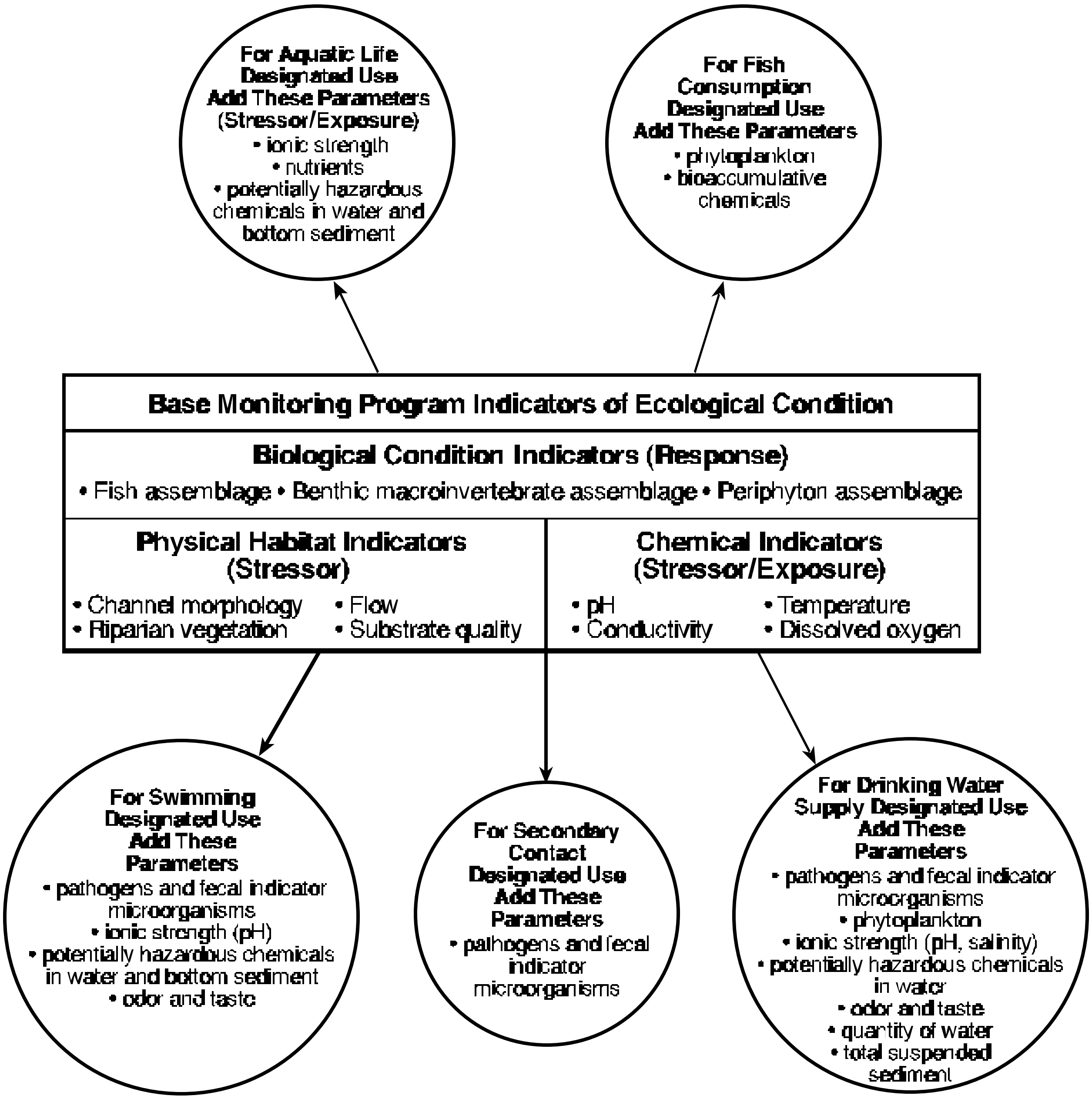

Figure 3-1. **ITFM Model for Stream Monitoring: Monitoring for different designated uses based on a combination of biological, physical, and chemical measures**

the parameters shown in Figure 3-1. Standard methods for measuring the chemical parameters and conducting toxicity tests are well established among the States, but methods for biological and habitat assessments are not standardized for all types of waterbodies. Recent work by the Ohio EPA suggests that bioassessment methods differ widely in their accuracy and discriminatory power for aquatic life use determinations (Yoder et al., 1994). Ohio evaluated a hierarchy of bioassessment approaches relevant to differing levels of rigor and confidence. In their State, Ohio EPA found that less intensive bioassessment approaches tend to be accurate **in detecting impairment**, but may give a false indication of full support in reaches where the methods are not rigorous enough to detect subtle problems.

ITFM (1995) recommends that to combine data for assessment, monitoring data produced by different organizations should be comparable, of known quality, available for integration with information from a variety of sources, and easily aggregated spatially and temporally. This is important at a variety of scales, up to and including national assessments. If different methods are similar with respect to the quality of data each produces, then data from those methods may be used interchangeably or together (Diamond et al. 1996). As data quality (i.e., precision, sensitivity) increases, the confidence in the assessment increases. Data quality objectives should be defined for each method so that assessments can be validated by imposing a known level of confidence in the results.

Monitoring Design

Any monitoring and assessment program begins with setting goals and a monitoring design that can meet those goals. The history of water quality monitoring is replete with programs that could not answer key questions. Examples include:

- A watershed study where the monitoring organization assumes that flow data can be obtained after the fact based on "reference point" measurements from bridges, only to learn later that many streams lack the channel morphometry to develop a stage-discharge relationship;

- An intensive survey where the laboratory's detection levels for metals prove inadequate to detect even concentrations above water quality standards;

- A basin survey where management or the legislature poses the question "What is the statistical trend in biological condition of our streams?" too late to be incorporated into the monitoring design.

As discussed in Section 2, EPA has a goal of comprehensively characterizing the Nation's streams, rivers, lakes, wetlands, estuaries, and shorelines. These assessments will include monitored and evaluated assessments and may involve probability-based as well as targeted monitoring. To achieve this goal, EPA encourages States to incorporate a formal process of goal setting and monitoring design while meeting their own State-specific goals. ITFM provides general guidelines for the topics to consider in monitoring design in a technical appendix of its final report (ITFM, 1995), and EPA's Section 106/604(b) monitoring guidance tailors the ITFM guidelines to the 106/305(b) process.

The Data Quality Objectives (DQO) process developed by EPA's Quality Assurance Management Staff is a specific approach to monitoring design that has been applied to monitoring programs in all media. The DQO process involves the stakeholders in the program in the design. Stakeholders itemize and clarify the questions being asked of a monitoring program, including the required level of accuracy in the answers. Generally, these questions are stated in quantitative terms ("What are the index of biotic integrity [IBI] and invertebrate community index [ICI] values for wadable streams in Big River Basin, and what is the trend in IBI across the basin, with 80 percent certainty?"), and statistical methods may be recommended for selecting sites or sampling frequency. For information about DQOs for water quality monitoring contact the Assessment and Watershed Protection Division at (202) 260-7023.

To date, States have taken three main approaches to monitoring a large portion of their waterbodies:

- Fixed-station networks with hundreds or thousands of sites (most large networks have been reduced in the past 10 years)

- Rotating basin surveys with a large number of monitoring sites covering thousands of miles of waters (Ohio EPA's bioassessment program)

- Rotating basin surveys with a probabilistic monitoring design; a statistically valid set of sites are selected for sampling in each basin (Delaware's benthic macroinvertebrate program).

The National Water Quality Monitoring Council may make recommendations about monitoring design; in the meantime, however, EPA encourages States to consider existing approaches such as Ohio's and Delaware's. In particular, EPA urges States to take advantage of monitoring data provided by other agencies such as USGS, NOAA, or the U.S. Fish and Wildlife Service (USFWS). See Section 2 for more information about comprehensive assessments using different monitoring designs.

3.2 Aquatic Life Use Support (ALUS)

The EPA/State 305(b) Consistency Workgroup has begun to implement the ITFM recommendations including how to integrate the results of biological, habitat, chemical and toxicological assessments in making a determination of aquatic life use support (ALUS). This approach includes consideration of assessment quality as indicated by levels of information of the different data types in evaluating the degree of impairment (partial support vs nonsupport) when there are differences in assessment results. Level of information is discussed below and described for each data type in Sections 3.2.1 through 3.2.4, Tables 3-1 through 3-4. Guidance on making assessments of ALUS for each individual data type is included in Sections 3.2.1 through 3.2.4. Guidance and case studies on integration of the assessment results from different data types, including consideration of level of information and site specific conditions, are presented in Section 3.2.5.

Level of Information

In 1994, the 305(b) Consistency Workgroup concluded that descriptive information characterizing the level of information, or rigor, in the method is needed to more fully define an assessment of use support. Documenting this information is important because users often need to know the basis of the underlying information. The Workgroup recommends that assessment quality information become a part of State assessment data bases. Consequently, the Workgroup has developed guidance for evaluating the level of information of methods used in making ALUS.

Data types are grouped into four categories: biological (Table 3-1), habitat (Table 3-2), toxicological (Table 3-3) and physical/chemical (Table 3-4). A hierarchy of methods corresponding to each data type and ordered by level of information is summarized in the tables. The rigor of a method within each data type is dictated by its technical components, spatial/temporal coverage, and data quality (precision and sensitivity). In the data type tables, Level 4 data are of highest quality for a data type and provide relatively high level of certainty. Level 1 data represent less rigorous approaches and thus provide a level of information with greater degree of uncertainty. However, in situations where severe conditions exist, a lower level of assessment quality will be adequate. For example, a severely degraded site can be characterized as impaired with a high level of confidence based on a cursory survey of biota or habitat, as in the case of repeated fish kills or severe sedimentation from mining. Data in Levels 1 through 4 vary in strengths and limitations, and, along with site-specific conditions, should be evaluated carefully for use in assessments. Data not adequate for ALUS determinations should be excluded from the assessment.

Table 3-1. Hierarchy of Bioassessment Approaches for Evaluation of Aquatic Life Use Attainment Based on Resident Assemblages

Level of Info[a]	Technical Components	Spatial/ Temporal Coverage	Data Quality[b]	WBS Codes[c]
1	Visual observation of biota; reference conditions not used; simple documentation	Limited monitoring; extrapolations from other sites	Unknown or low precision and sensitivity; professional biologist not required	310, 320, 350, 322
2	One assemblage (usually invertebrates); reference conditions pre-established by professional biologist; biotic index or narrative evaluation of historical records	Limited to a single sampling; limited sampling for site-specific studies	Low to moderate precision and sensitivity; professional biologist may provide oversight	310, 320, 322, 350
3	Single assemblage usually the norm; reference condition may be site-specific, or composite of sites (e.g., regional); biotic index (interpretation may be supplemented by narrative evaluation of historical records)	Monitoring of targeted sites during a single season; may be limited sampling for site-specific studies; may include limited spatial coverage for watershed-level assessments	Moderate precision and sensitivity; professional biologist performs survey or provides training for sampling; professional biologist performs assessment.	310, 315, 320, 321, 330, 331, 350
4	Generally two assemblages, but may be one if high data quality; regional (usually based on sites) reference conditions used; biotic index (single dimension or multimetric index)	Monitoring during 1-2 sampling seasons; broad coverage of sites for either site-specific or watershed assessments; conducive to regional assessments using targeted or probabilistic design	High precision and sensitivity; professional biologist performs survey and assessment	310, 315, 320, 321, 330, 331, 340, 350

NOTE: Table is based on use in lotic systems. With some modification, these approaches would apply to other waterbody types.

[a] Level of information refers to rigor of bioassessment, where 1 = lowest and 4 = highest.

[b] Refers to ability of the ecological endpoints to detect impairment or to differentiate along a gradient of environmental conditions.

[c] WBS Assessment Type Codes from Table 1-1.

Table 3-2. Hierarchy of Habitat Assessment Approaches for Evaluation of Aquatic Life Use Attainment

Level Of Info[a]	Technical Components	Spatial/ Temporal Coverage	Data Quality[b]	WBS Codes[c]
1	Visual observation of habitat characteristics; no true assessment; documentation of readily discernable land use characteristics that might alter habitat quality; no reference conditions	Sporadic visits; sites are mostly from road crossings or other easy access	Unknown or low precision and sensitivity; professional scientist (biologist, hydrologist) not required	365
2	Visual observation of habitat characteristics and simple assessment; use of land use maps for characterizing watershed condition; reference condition pre-established by professional scientist	Limited to annual visits and non-specific to season; generally easy access; limited spatial coverage and/or site-specific studies	Low precision and sensitivity; professional biologist or hydrologist not involved or only correspondence	370
3	Visual-based habitat assessment using standard operating procedures (SOPs); may be supplemented with quantitative measurements of selected parameters; conducted with bioassessment; data on land use compiled and used to supplement assessment; reference condition used as a basis for assessment	Assessment during a single season usually the norm; spatial coverage may be limited or broad and commensurate with biological sampling; assessment may be regional or site-specific	Moderate precision and sensitivity; professional biologist or hydrologist performs survey or provides oversight and training	375
4	Assessment of habitat based on quantitative measurements of instream parameters, channel morphology, and floodplain characteristics; conducted with bioassessment; data on land use compiled and used to supplement assessment; reference condition used as a basis for assessment	Assessment during 1-2 seasons; spatial coverage usually broad and commensurate with biological sampling; assessment may be regional or site-specific	High precision and sensitivity; professional biologist or hydrologist performs survey and assessment	380

NOTE: Table is based on use in lotic systems. With some modification, these approaches would apply to other waterbody types.

[a] Level of information refers to rigor of habitat assessment, where 1 = lowest and 4 = highest.

[b] Refers to ability of the habitat endpoints to detect impairment or to differentiate along a gradient of environmental conditions.

[c] WBS Assessment Type Codes from Table 1-1.

Table 3-3. Hierarchy of Toxicological Approaches and Levels for Evaluation of Aquatic Life Use Attainment

Level of Info[a]	Technical Components	Spatial/ Temporal Coverage	Data Quality[b]	WBS Codes[c]
1	Any <u>one</u> of the following: C Acute or chronic WET C Acute ambient C Acute sediment	1-2 WET tests/yr or 1 ambient or sediment sample tested in a segment or site	Unknown/low; minimal replication used; laboratory quality or expertise unknown	510, 520, 530, 550
2	Any of the following: C Acute <u>or</u> chronic ambient C Acute sediment C Acute <u>and</u> chronic WET for effluent-dominated system	3-4 WET tests/yr or 2 ambient or sediment samples tested in a segment or site at different times	Low/moderate—little replication used within a site; laboratory quality or expertise unknown or low	510, 520, 530, 540, 550
3	Any of the following: C Acute and chronic WET for effluent-dominated system C Chronic ambient <u>or</u> acute or chronic sediment	Monthly WET tests or total of 3 tests based on samples collected in a segment at 3 different times	Moderate/high—replication used; trained personnel and good laboratory quality	510, 520, 540, 550
4	Both of the following: C Acute and chronic ambient and C Acute <u>or</u> chronic sediment	$ 4 tests in total based on samples collected in a segment at 4 different times including low flow conditions	High—replication used; trained personnel and good laboratory quality	530, 540, 550

[a] Level of information refers to rigor of toxicity testing, where 1 = lowest and 4 = highest

[b] Refers to ability of the toxicity testing endpoints to detect impairment or to differentiate along a gradient of environmental conditions

[c] WBS Assessment Type Codes from Table 1-1.

Table 3-4. Hierarchy of Physical/chemical Data Levels for Evaluation of Aquatic Life Use Attainment

Level of Info[a]	Technical Components	Spatial/Temporal Coverage	Data Quality[c]	WBS Codes[d]
1	Any <u>one</u> of the following: C Water quality monitoring using grab water sampling C Water data extrapolated from an upstream or downstream station where homogeneous conditions are expected C Monitoring data >5 years old without further validation C Best professional judgment based on land use data, source locations	Low spatial and temporal coverage: C Quarterly or less frequent sampling with limited period of record (e.g., 1 day) C Limited data during key periods or at high or low flows (critical hydrological regimes)[b].	Unknown/ Low	210, 220, 230, 240, 850, 150, 130
2	Any <u>one</u> of the following: C Water quality monitoring using grab water sampling C Rotating basin surveys involving multiple visits or automatic sampling C Synthesis of existing or historical information on fish contamination levels C Screening models based on loadings data (not calibrated or verified).	Moderate spatial and temporal coverage: C Bimonthly or quarterly sampling during key periods (e.g., spring/ summer months C Fish spawning seasons, including limited water quality data at high and low flows C Short period of record over a period of days or multiple visits during a year or season.	Low/ Moderate	210, 220, 222, 230, 240, 242, 260, 810, 180
3	Any <u>one</u> of the following: C Composite or a series of grab water sampling used (diurnal coverage as appropriate) C Calibrated models (calibration data <5 years old).	Broad spatial and temporal (long-term, e.g., > 3 years) coverage of site with sufficient frequency and coverage to capture acute events: C Typically, monthly sampling during key periods (e.g., spring/ summer months, fish spawning seasons), multiple samples at high and low flows C Lengthy period of record (sampling over a period of months).	Moderate/ High	211, 222, 242, 250, 610
4	All of the following: C Water quality monitoring using composite or series or grab samples (diurnal coverage as appropriate) C Limited sediment quality sampling and fish tissue analyses at sites with high probability of contamination.	Broad spatial (several sites) and temporal (long-term, e.g., > 3 years) coverage of site with sufficient frequency and parametric coverage to capture acute events, chronic conditions, and all other potential P/C impacts C Monthly sampling during key periods (e.g., spring/summer months C Fish spawning seasons) including multiple samples at high and low flows C Continuous monitoring.	High	231, 242, 250

NOTE: Physical refers to physical water parameters (e.g., temperature, pH, dissolved oxygen, turbidity, color, conductivity)

[a] Level of information refers to rigor of physical/chemical sampling and analysis, where 1 = lowest and 4 = highest.

[b] Even a short period of record can indicate a high confidence of *impairment* based on P/C data; 3 years of data are not required to demonstrate impairment. For example, a single visit to a stream with severe acid mine drainage impacts (high metals, low pH) can result in high confidence of nonsupport. However, long-term

monitoring may be needed to establish full support.
[c] Refers to ability of the physical/chemical endpoints to detect impairment or to differentiate along a gradient of environmental conditions.
[d] WBS Assessment Type Codes from Table 1-1.

At the Workgroup's recommendation, EPA is applying levels of information to wadable streams and rivers where EPA's Rapid Bioassessment Protocols or other comparable methods can be applied. This is because, at this time, monitoring methods for wadable streams and rivers are better documented and standardized (Gibson et al. 1996, Plafkin et al., 1989) than for other surface water resources such as lakes and estuaries.

EPA asks States to document the level of information that characterizes their methods for biological, habitat, toxicological, and chemical evaluations. The approach may be extended to ALUS determinations in other types of waterbodies as well as other designated uses in future 305(b) cycles based on the experience with ALUS in streams and rivers and as methods for other waterbody types are standardized. The Waterbody System will contain fields to track level of information for each data type (first columns of Tables 3-1 through 3-4).

EPA encourages States to store and provide this information for each river and stream assessment in addition to WBS Assessment Type Codes. See Section 6, especially Table 6-1, of the main *Guidelines* volume regarding data elements for annual electronic reporting.

3.2.1 Bioassessment

Biological survey methods are desirable for ALUS determinations, because they measure ecosystem health and integrity more directly than surrogate techniques and serve as response indicators to a variety of stressors. Certain biological survey and assessment techniques are useful for screening; i.e., they are intended to be sufficient for detecting problems and may not be as rigorous as techniques used to assess the degree of use support or prioritize sites for further study or some mitigation action. However, simple biological screening techniques are usually sufficient to identify severely degraded or the other extreme (i.e., excellent) biological conditions. A hierarchy of biological approaches can be developed that incorporates certain technical considerations and are relevant to various levels of information (Table 3-1). The data quality elements emphasize a determination of precision (i.e., measurement error at a site as evidenced by the reproducibility of metric values or bioassessment scores for a given site during the same index period) and sensitivity (i.e., the ability to detect impairment relative to the reference condition).

Based on considerable information already available, EPA strongly endorses the regional reference approach for State bioassessment programs for streams (Gibson et al. 1996), which is a level 3 or 4 assessment in Table 3-1. If States choose not to implement a reference site approach, they are still encouraged to monitor two organism assemblages (level 4), with detailed taxonomy, a multimetric approach, and habitat evaluation. In calling for two assemblages, EPA seeks to include critical groups in the food chain that may react to different ecosystem stressors or differently to the same stressor. EPA recognizes that the use of two assemblages or the regional reference approach may not be feasible in certain cases (e.g., streams in the arid west due to naturally occurring conditions such as extreme temperatures and lack of flow). EPA also recognizes that some State bioassessment programs are in their early stages and may not yet have the capability to use a regional reference site approach or to monitor more than one assemblage.

Many States (Davis et al. 1996) are currently assessing a single assemblage, benthic macroinvertebrates, with detailed taxonomy, a multimetric approach, and habitat evaluation (Level 2 or 3 assessment in Table 3-1). These States are monitoring a critical assemblage that often gives the greatest information about ecosystem health for the available resources. For fish sampling, some rely on their fish and game agencies, which are mainly oriented to game fish. As resources permit, EPA encourages State water quality agencies to develop the capability for fish assemblage monitoring themselves or work with the fish and game staff to develop the needed capabilities.

<u>ALUS Determination Based on Biological Assessment Data</u>

A. Fully Supporting: Reliable data indicate functioning, sustainable biological assemblages (e.g., fish, macroinvertebrates, or algae) none of which has been modified significantly beyond the natural range of the reference condition.

B. Partially Supporting: At least one assemblage (e.g., fish, macroinvertebrates, or algae) indicates moderate modification of the biological community compared to the reference condition.

C. Not Supporting: At least one assemblage indicates nonsupport. Data clearly indicate severe modification of the biological community compared to the reference condition.

The interpretation of the terms "modified significantly," "moderate modification," and "severe modification" is State-specific and depends on the State's monitoring and water quality standards programs. For example, Ohio EPA reports nonattainment (not supporting) if none of its 3 indices (2 for fish and 1 for macroinvertebrates) meet ecoregion criteria or if one assemblage indicates severe toxic impact (Ohio's poor or very poor category), even if the other assemblage indicates attainment. Partial support exists if 1 of 2 or 2 of 3 indices do not meet ecoregion criteria and are in the poor or very poor category.

Additional Considerations for Lakes

State lake managers should address more than one biological assemblage in making lake ALUS decisions. Many parameters of these assemblages may not have specific criteria (e.g., algal blooms, growth of nuisance weeds) but have important effects on lake uses. Many are also response indicators of the level of lake eutrophication.

Lake resources vary regionally, even within States, due to variations in geology, vegetation, hydrology, and land use. Therefore, regional patterns of lake water quality, morphometry (physical characteristics such as size, shape, and depth), and watershed characteristics should ideally be defined based on comparison to natural conditions using an ecoregion approach. The State can then set reasonable goals and criteria for a variety of parameters. These regional patterns currently apply to natural lakes, but are being evaluated for use with reservoirs.

EPA is developing guidance on bioassessment protocols and biological criteria development for lakes and reservoirs (*Guidance on Lake and Reservoir Bioassessment and Biocriteria*, draft, U.S. EPA, 1996). Draft guidance is currently being revised to address a review of comments by EPA's Science Advisory Board. Notice of availability for public review and comment in the *Federal Register* is planned for 1997.

3.2.2. Habitat Assessment

Assessment of the physical habitat structure is necessary for aquatic life support evaluations because the condition and/or potential of the biological community is dependent upon supportive habitat. Aquatic fauna often have very specific habitat requirements, independent of water quality (Barbour et al. 1996a). The technique of habitat assessment has evolved substantially over the last decade to provide adequate information on the quality of the habitat. Numerous State and Tribal agencies are well-versed in habitat assessment and have incorporated appropriate techniques into their monitoring programs. Results from nonpoint-source assessments suggest that habitat alteration is a major source of perturbation of the Nation’s surface waters. The strengths of habitat assessment are: (1) enhances interpretation of biological data; (2) provides information on non-chemical stressors, and (3) leads to informed decisions regarding problem identification and restoration.

Most often, habitat assessment is conducted in conjunction with bioassessment. A general habitat assessment incorporates physical attributes from microhabitat features such as substrate, velocity, depth, to channel morphology features such as width, sinuosity, flow or volume, to riparian and bank structure features. All of these features are stressor indicators. The approach also can integrate habitat information into an index or summary of overall habitat condition.

The rigor of the habitat assessment ranges from a visual-based characterization (Level 1), which documents specific characteristics without placing a value, to a true assessment (Levels 2 through 4), which places a value on the quality of the physical habitat structure (Table 3-2). Habitat assessments may be visual-based (e.g., RBPs), patterned after Ohio EPA (1987), Plafkin et al. (1989), Florida DEP (1994), and Idaho DEQ (1995), or more quantitative as suggested by the Environmental Monitoring and Assessment Program (EMAP). The data quality associated with habitat assessment is more difficult to define than with bioassessment, but can be done by a comparison among investigators.

<u>ALUS Determination Based on Habitat Assessment Data</u>.

A. Fully Supporting: Reliable data indicate natural channel morphology, substrate composition, bank/riparian structure, and flow regime of region. Riparian vegetation of natural types and of relatively full standing crop biomass (i.e., minimal grazing or disruptive pressure).

B. Partially Supporting: Modification of habitat slight to moderate usually due to road crossings, limited riparian zones because of encroaching land use patterns, and some watershed erosion. Channel modification slight to moderate.

C. Not Supporting: Moderate to severe habitat alteration by channelization and dredging activities, removal of riparian vegetation, bank failure, heavy watershed erosion or alteration of flow regime.

Habitat assessment is mostly conducted in conjunction with bioassessment. However, degradation of habitat associated with aquatic resources is a primary stressor limiting the attainment of aquatic life use support in many regions of the country. Land use patterns involving urban development and impervious surface, agriculture and ranching, silviculture, mining, and flood control/regulation are generally the principal factors in habitat degradation.

3.2.3. Aquatic and Sediment Toxicity Methods

EPA recommends that information from toxicity tests be separated from the physical/chemical data. Although chemical criteria are based on toxicity tests, actual testing done to evaluate an aquatic life use should be treated as an additional ecological indicator.

Toxicity tests are a well-established tool for examining effects of both point and nonpoint sources of chemicals or effluents in surface waters (i.e., stressor and exposure indicators). Most States require whole effluent toxicity (WET) testing of waste water dischargers under the NPDES program. For ALUS, ambient water column and whole sediment toxicity tests may be most relevant, particularly if the early life stages of test organisms and sublethal (chronic) endpoints are used (Table 3-3). Ambient tests use samples that are collected from sites and that are typically used whole (i.e., no dilution). Toxicity tests, like chemical analyses, use temporally discrete samples which, in the case of water column tests, typically have short holding times (< 36 hours according to EPA guidance). Sediment samples may be held for longer periods (2 to 8 weeks) prior to testing if stored properly. Samples used in aquatic toxicity testing are usually collected over no more than a 24-hour period. Sediment samples, by their very nature, are grab samples which are also collected over a short time period (hours) at any one site. As a result, all toxicity tests, even those involving prolonged chronic exposures (such as EPA 7-day chronic tests or 28-day chronic sediment tests), yield data that are a “snapshot” in time. The longer the period of time over which site water or sediment samples are collected and used in testing, the longer the “snapshot” and the higher confidence that the test result is representative of prevailing water or sediment quality conditions at that time. The strengths of ambient toxicity tests are:

- They aid in identifying point and nonpoint source water-quality impairments that may otherwise be undetectable using other monitoring tools;

- They are used for confirming that observed impairment is not due to chemical or toxicity-related sources. Ohio EPA and the North Carolina Division of Water Quality, for example, used toxicity tests to demonstrate that habitat or physical stressors were the major causes of impairment in some systems and not point-source toxicity as previously assumed;

- They integrate biological effects of most chemical stressors present, thereby giving a more accurate estimate of the actual water or sediment quality as compared to chemical concentration measurements; this has been shown to be particularly true for certain water column metals, bulk sediment chemical measurements that do not take into account total organic carbon or acid volatile sulfide concentrations (for nonpolar organics and metals, respectively), and for sites in which potential pollutants were unmeasured or unknown.

WET tests are potentially useful for ALUS at sites in which an effluent contributes the major flow instream (i.e., effluent-dominated or effluent-dependent systems). These tests are well standardized and relatively easy to interpret, however, their relationship to ALUS is dependent on many factors that may or may not be identifiable for the system of interest (Waller et al. 1996; LaPoint et al. 1996).

Sediment toxicity tests are especially useful for ALUS since sediments can be prominent sources as well as sinks. For this reason, sediment samples may represent a somewhat longer "snapshot" in time than water column samples. Also, because sediment samples can be stored for longer periods than water samples, they are more convenient to use in testing. Collection of sediment pore water or elutriates further enhances the use of sediments in ALUS because these fractions may contain most of the bioavailable pollutants present and because these fractions are amenable to standard aquatic toxicity test methods. Combined with bioassessments and sediment chemical analyses, sediment toxicity is a powerful tool to evaluate and identify causes of impairment. Whole sediment testing, using the more standardized 10-day acute tests, may be most appropriate for ALUS. These are the least labor-intensive and costly tests and are also easiest to interpret. The more recently developed EPA chronic sediment test methods (which should be available by the end of 1997) are also promising tools for ALUS. Sediment testing is most relevant if there are appropriate reference site sediments available with which to compare different site samples. Usually, such reference sites are available, but in some instances are defined by trial and error. The use of clean laboratory-formulated reference sediments as a means of comparison is also a viable option, particularly if factors such as sediment particle size are similar to that observed at the site of interest.

Concerns with sediment tests are: (1) for representativeness, many sediment samples may need to be composited at a site to overcome physical and chemical heterogeneity; (2) storage and manipulation of samples prior to testing may change the chemical characteristics and toxicity of a sample in unknown ways; and (3) for some species, physical characteristics of the sediment (e.g., particle size or TOC) may be suboptimal for the test species resulting in a false positive or apparently toxic conditions when there are none. This may necessitate the use of two or more different test species for a given sediment sample.

Several EPA, American Society for Testing Materials (ASTM), and State agency toxicity test methods exist, both for saltwater and freshwater aquatic and sediment toxicity tests, ranging from short-term acute or lethality tests (usually 48 to 96h in length for aquatic and pore water or elutriate tests and 10d for whole sediments) to longer term early life stage (7 day for pore water and elutriates and 28 day for whole sediments) and full life-cycle (> 21 day for aquatic tests) chronic tests that measure sublethal endpoints. Some sublethal tests such as those for saltwater bivalve embryo-larval development or echinoderm fertilization, may be much shorter in duration (48 and 1.5 hour, respectively). Appropriate sample collection is critical to ensure representative and accurate results. In addition, chemically inert sampling equipment must be used and depth and/or width integrated composite samples should be considered for ALUS determination.

ALUS Determinations Based on Aquatic and/or Sediment Toxicity Data

A. Fully Supporting: No toxicity noted in either acute or chronic tests compared to controls or reference conditions.

B. Partially Supporting: No toxicity noted in acute tests, but may be present in chronic tests in either slight amounts and/or infrequently within an annual cycle.

C. Not Supporting: Toxicity noted in many tests and occurs frequently.

Other Considerations

For certain species such as planktonic ones, ambient aquatic samples may appear more or less toxic due to the presence of certain natural water quality conditions or eutrophication effects. Ambient tests are a “snapshot” in time and may be unrepresentative of other times, seasons, or flows. Non-toxic conditions include naturally high dissolved solids, hardness, or conductivity, or naturally low alkalinity and hardness. Appropriate reference site or control samples for comparison may not be readily available in some systems resulting in a certain amount of uncertainty in extrapolating laboratory control or simulated reference conditions to actual natural conditions at a site. WET tests are best incorporated into the NPDES program; for ALUS, the results obtained using tools in the 305(b) process such as bioassessment, ambient aquatic and sediment toxicity tests, and chemical monitoring are more appropriate.

3.2.4 Physical/Chemical Methods

The use of physical/chemical data as stressor and exposure indicators for determining ALUS has long been a basis of State monitoring programs. Established criteria exist for many chemical parameters and standard sampling and analysis protocols have been developed for ensuring consistency and quality control. These data are separated into categories of toxicants (priority pollutants, chlorine, and ammonia), conventionals (dissolved oxygen, pH, temperature) in reference to the physical constituents of water quality, and metals. Although SOPs exist for physical/chemical parameters, States still differ in their design and implementation of chemical sampling and analysis (Table 3-4). Sampling frequency and intensity vary among states. The number of parameters sampled and analyzed also varies among programs which influences comparability in assessments.

Analyses of chemical concentrations in fish tissues are included in Table 3-4. Though not a traditional or required measure of ALUS, fish tissue concentrations are useful for evaluating the potential impacts to wildlife that depend on aquatic systems for food and/or habitat.

ALUS Determinations Based on Physical/Chemical Assessment Data

EPA recognizes that many States may not always collect a broad spectrum of chemical data for every waterbody. Therefore, States are expected to apply the following guidance to whatever data are available and to use a "worst case" approach where multiple types of data are available. If, for example, chemical data indicate full support but temperature data indicate impairment, the waterbody is considered impaired.

Conventionals (dissolved oxygen, pH, temperature)

A. Fully Supporting: For any one pollutant or stressor, criteria exceeded in #10 percent of measurements. In the case of dissolved oxygen (DO), national ambient water quality criteria specify the recommended acceptable daily average and 7-day average minimums and the acceptable 7-day and 30-day averages. States should document the DO criteria being used for the assessment and should discuss any biases that may be introduced by the sampling program (e.g., grab sampling in waterbodies with considerable diurnal variation).

B. Partially Supporting: For any one pollutant, criteria exceeded in 11 to 25 percent of measurements. For DO, the above considerations apply.

C. Not Supporting: For any one pollutant, criteria exceeded in >25 percent of measurements. For DO, the above considerations apply.

Special Considerations for Lakes

For lakes, States should discuss their interpretation of DO, pH, and temperature standards for both epilimnetic and hypolimnetic waters. In addition, States should consider turbidity and lake bottom siltation.

Toxicants (priority pollutants, metals, chlorine, and ammonia)

A. Fully Supporting: For any one pollutant, no more than 1 exceedance of acute criteria (EPA's criteria maximum concentration or applicable State/Tribal criteria) within a 3-year period based on grab or composite samples and no more than 1 exceedance of chronic criteria (EPA's criteria continuous concentration or applicable State/Tribal criteria) within a 3-year period based on grab or composite samples.

B. Partially Supporting: For any one pollutant, acute or chronic criteria exceeded more than once within a 3-year period, but in ≤ 10 percent of samples.

C. Not Supporting: For any one pollutant, acute or chronic criteria exceeded in >10 percent of samples.

Note: The above assumes at least 10 samples over a 3-year period. If fewer than 10 samples are available, the State should use discretion and consider other factors such as the number of pollutants having a single violation and the magnitude of the exceedance(s).

Other Considerations Regarding Toxicant Data

- EPA maintains that chronic criteria should be met in a waterbody that fully supports its uses. Few States and Tribes, if any, are obtaining composite data over a 4-day sampling period for comparison to chronic criteria. EPA believes that 4-day composites are not an absolute requirement for evaluating whether chronic criteria are being met. Grab and composite samples (including 1-day composites) can be used in water quality assessments if taken during stable conditions.

This should give States more flexibility in utilizing chronic criteria for assessments.

- States should document their sampling frequency. Sampling frequency should be based on potential variability in toxicant concentrations. In general, waters should have at least quarterly data to be considered monitored; monthly or more frequent data are considered abundant. More than 3 years of data may be used, although the once-in-3-years consideration still applies (i.e., two violations are allowed in 6 years of abundant data).

- The once-in-3-years goal is not intended to include spurious violations resulting from lack of precision in analytical tests. Therefore, using documented quality assurance/quality control (QA/QC) assessments, States may consider the effect of laboratory imprecision on the observed frequency of violations.

- If the duration and frequency specifications of EPA criteria change in the future, these recommendations should be changed accordingly.

- Samples should be taken outside of designated mixing zones or zones of initial dilution.

Special Considerations Regarding Metals

The implementation and application of metals criteria is complex due to the site-specific nature of metals toxicity. EPA's policy is for States to adopt and use the dissolved metal fraction to set and measure compliance with water quality standards, because dissolved metal more closely approximates the bioavailable fraction of metal in the water column than does total recoverable metal. One reason is that a primary mechanism for water column toxicity is adsorption at the gill surface which requires metals to be in the dissolved form. Table 3-5 provides guidance for calculating EPA dissolved criteria from the published total recoverable criteria. The dissolved metal criteria, expressed as percentage, are presented as recommended values and ranges. If a State is collecting dissolved metal data but does not yet have dissolved criteria, Table 3-5 might be useful for estimating screening values. Also, if total recoverable metal concentrations are less than the estimated dissolved metal criteria calculated from Table 3-5, the State could be relatively certain that toxic concentrations are not present.

Some States have already developed and are using dissolved metals criteria and should continue to do so. In the absence of dissolved metals data and State criteria, States should continue to apply total recoverable metals criteria to total recoverable metals data because this is more conservative and thus protective of aquatic life. In some situations, a State may choose to use total recoverable metals criteria when there are indications that total metal loadings could be a stress to the ecosystem. The ambient water quality criteria are neither designed nor intended to address the fate and effect of metals in an ecosystem, e.g., protect sediments, prevent effects due to food webs containing organisms that dwell in the sediments and those that dwell in the water column and filter or ingest suspended particles. However, since consideration of sediments or bioaccumulative impacts is not incorporated into the criteria methodology, the appropriateness and degree of conservatism inherent in the total recoverable approach is unknown.

Historical metals data should be used with care. Concern about the reliability of the data are greatest below about 5 to 10 ppb due to the possibility of contamination problems during sample collection and analysis. EPA believes that most historical metals concentrations above this level are valid if collected with appropriate quality assurance and quality control.

Table 3-5. Recommended Factors for Converting Total Recoverable Metal Criteria to Dissolved Metal Criteria

Metal	Recommended Conversion Factors	
	CMC[a]	CCC[a]
Arsenic (III)	1.000	1.000
Cadmium[b]		
Hardness = 50 mg/L	0.973	0.938
Hardness = 100 mg/L	0.944	0.909
Hardness = 200 mg/L	0.915	0.880
Chromium (III)	0.316	0.860[c]
Chromium (VI)	0.982	0.962
Copper	0.960	0.960
Lead[b]		
Hardness = 50 mg/L	0.892	0.892
Hardness = 100 mg/L	0.791	0.791
Hardness = 200 mg/L	0.690	0.690
Nickel	0.998	0.997
Selenium	0.922	0.922
Zinc	0.978	0.986

[a] CMC = Criterion Maximum Concentration
CCC = Criterion Continuous Concentration

[b] The recommended conversion factors (CFs) for any hardness can be calculated using the following equations:

Cadmium
CMC: CF = 1.136672 - [(ln hardness) (0.041838)]
CCC: CF = 1.101672 - [(ln hardness) (0.041838)]

Lead (CMC and CCC): CF = 1.46203 - [(ln hardness) (0.145712)]

where:
(ln hardness) = natural logarithm of the hardness. The recommended CFs are given to three decimal places because they are intermediate values in the calculation of dissolved criteria.

[c] This CF applies only if the CCC is based on the test by Stevens and Chapman (1984). If the CCC is based on other chronic tests, it is likely that the CF should be 0.590, 0.376, or the average of these two values.

Source: Stephen, C. E. 1995. *Derivation of Conversion Factors for the Calculation of Dissolved Freshwater Aquatic Life Criteria for Metals*. U.S. EPA, Environmental Research Laboratory, Duluth.

3.2.5 Integration of Different Data Types in Making an ALUS Determination

The following guidelines apply to ALUS determinations for wadable streams and rivers when biological, habitat, chemical, and/or toxicity data types are available (Figure 3-2, Table 3-6). These guidelines strongly emphasize the use of biological data for the assessment of ALUS specific to wadeable streams and rivers. However, the basic principles are applicable to other waterbody types. This guidance has undergone external peer-review (Dickson et al. 1996) and has been revised to address the principle peer-review recommendations to improve the guidance. In addition, peer review recommendations were made to expand the guidance to (1) develop a confidence icon for the overall assessment and (2) develop guidelines that consider the results from biological, chemical and physical assessments in relation to their role as response, stressor or exposure indicators. The peer review specifically recommended that EPA develop a weighting algorithm for biological results (as response indicator) in relation to results from physical/chemical, habitat, and toxicological assessments (as stressor/exposure indicators). These latter recommendations will be evaluated for future guidelines. EPA considers the current guidelines, particularly consideration of level of information, as providing the initial basis for addressing these additional peer review recommendations.

EPA recommends consideration of the level of information of the different data types in evaluating degree of impairment (partial support vs nonsupport). Case studies follow that demonstrate how ALUS determinations could be made based on types of data, level of information, and site specific information and conditions, and are not intended to cover all possible situations but to highlight commonly encountered scenarios. These case studies are based on actual State examples that represent a State's decision process in making an ALUS determination, and are presented in a uniform manner for illustration. Different states use different ordinal scales for assessment.

Generally, assessments based on data with high levels of information should be weighted more heavily than those based on data with low levels of information, and biological data should be weighted more heavily than other data types. In particular, it is recommended that the results of biological assessments, especially those with high levels of information, be the basis for the overall ALUS determination if the data indicate impairment. This is because the biological data provide a direct measure of the status of the aquatic biota and detect the cumulative impact of multiple stressors on the aquatic community, including new or previously undetected stressors. This approach is consistent with EPA's

Policy on Independent Application while incorporating a weight of evidence approach in determining the degree of impairment (partial or nonsupport). The Policy does not allow for a

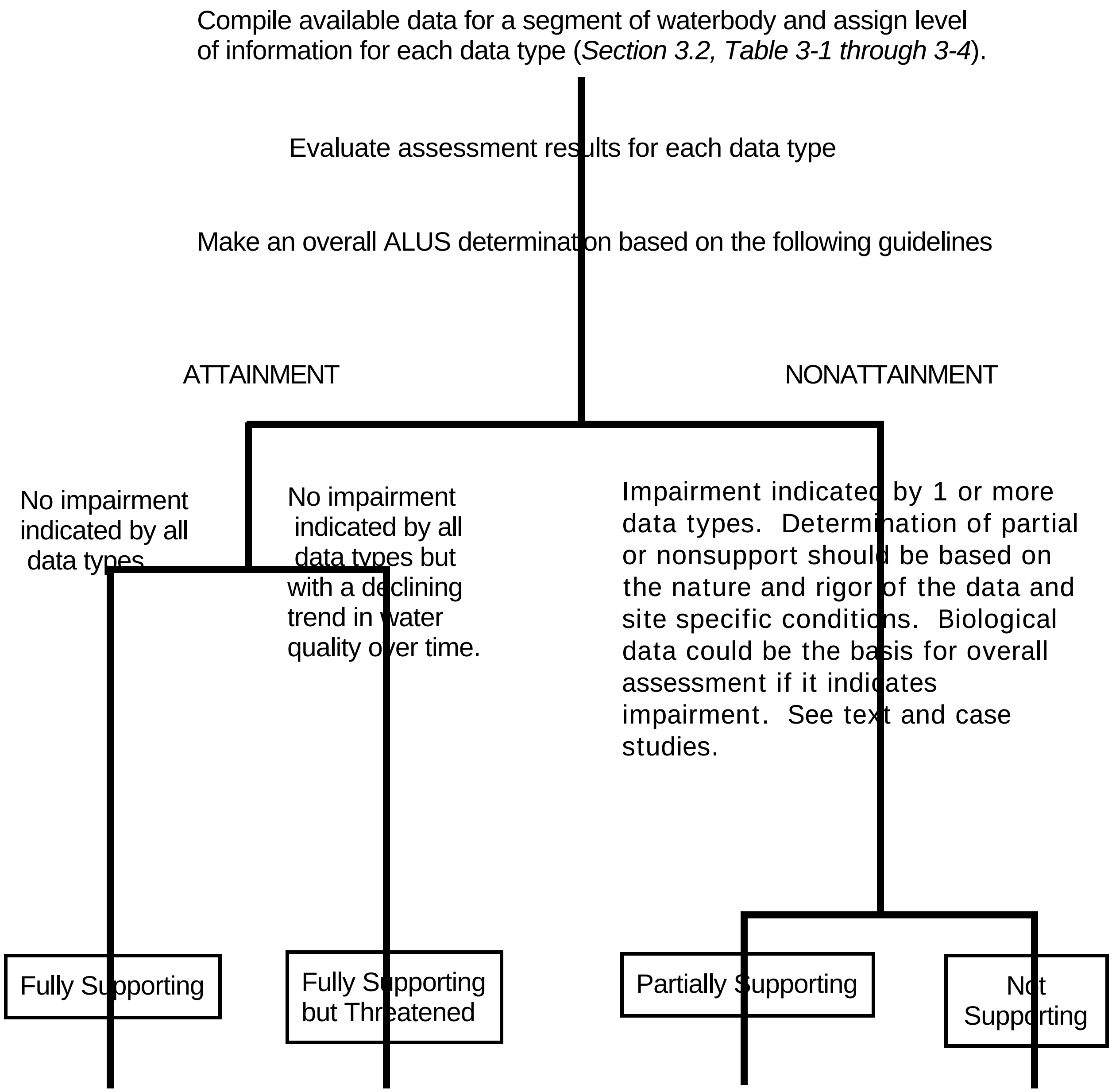

Figure 3-2. Determination of ALUS using biological, chemical, toxicological, and/or habitat data.

Table 3-6. Determination of ALUS Using More Than One Data Type

ALUS Attainment	
A. Fully Supporting:	No impairment indicated by all data types
B. Fully Supporting but Threatened:	No impairment indicated by all data types; one or more categories indicate an apparent decline in ecological quality over time or potential water quality problems requiring additional data or verification, or Other information suggests a threatened determination (see Section 3.2)
ALUS Non-attainment	
C. *Partially Supporting:	Impairment indicated by one or more data types and no impairment indicated by others.
D. *Not Supporting:	Impairment indicated by all data types
* A determination of *partially supporting* or *not supporting* could be made based on the nature and rigor of the data and site-specific conditions in the results of the data types. If bioassessment (usually Level 3 or 4) indicates impairment, then a determination of not supporting should be made. See case studies that follow.	

Case studies (3 pages)--see separate file “V2, CHAP 3--CASE STUDIES”

determination of full support when there are differences in assessment results when at least one assessment indicates impairment. For example, it is possible to arrive at an overall assessment of partial support where biological data indicate full support and other data types indicate some level of impairment.

3.2.6 Additional Information on Biological Assessment of ALUS for Wadable Streams and Rivers

The following information may be useful to States in making ALUS determinations based on biological and associated habitat data. Biological assessments are evaluations of the biological condition of waterbodies using biological surveys and other direct measurements of resident biota in surface waters and comparing results to the established biological criteria. They are done by qualified professional staff trained in biological methods and data interpretation. The utility of biological measures has been demonstrated in assessing impairment of receiving waterbodies, particularly that caused by nonpoint sources and nontraditional water quality problems such as habitat degradation. Biological assessments are key to determining whether functional, sustainable communities are present and whether any of these communities have been modified beyond the natural range of the reference condition. Functional and sustainable implies that communities at each trophic level have species composition, population density, tolerance to stressors, and healthy individuals within the range of the reference condition and that the entire aquatic system is capable of maintaining its levels of diversity and natural processes in the future (see Angermeier and Karr, 1994).

The techniques for biosurveys are still evolving, but there have been significant improvements in the last decade. Appropriate methods have been established by EPA (e.g., Plafkin et al., 1989), State agencies (e.g., Ohio EPA, 1987; Massachusetts DEP, 1996; Florida DEP, 1994; Idaho DEQ, 1995), and other investigators assessing the condition of the biota (e.g., Karr et al., 1986). Guidance for development of biocriteria-based programs is provided in the *Biological Criteria: National Program Guidance for Surface Waters* (U.S. EPA, 1990) and *Biological Criteria: Technical Guidance for Streams and Small Rivers* (Gibson et al., 1996). As biosurvey techniques continue to improve, several technical considerations apply:

- *The identification of the REFERENCE CONDITION is basic to any assessment of impairment or attainment of aquatic life use and to the establishment of biological criteria.*

Reference conditions are described from an aggregate of data best acquired from multiple sites with similar physical dimensions, represent minimally impaired conditions, and provide an estimate of natural variability in biological condition and habitat quality. For determining reference condition, alternative approaches to selection of reference sites include use of historical data, paleoecological data for lakes, experimental laboratory data for select cases, quantitative models, and best professional judgment (Hughes 1995).

Reference conditions must be stratified (i.e., put into homogenous waterbody classes) to account for much of the natural physical and climatic variability that affects the geographic distribution of biological communities. The Ecoregion Concept (Omernik, 1987) recognizes geographic patterns of similarity among ecosystems, grouped on the basis of environmental variables such as climate, soil type, physiography, and vegetation. Currently, efforts are under way in several parts of the country to refine these ecoregions into a more useful framework to classify waterbodies. Procedures have begun in several ecoregions and subecoregions to identify reference conditions within those particular units. In essence, these studies are developing reference databases to define biological potential and physical habitat expectations within ecoregions. The concept of reference conditions for bioassessment and biocriteria is discussed further below.

In developing community bioassessment protocols, reference conditions against which to compare test sites and to judge impairment are needed. Ideally, reference conditions represent the highest biological conditions found in waterbodies unimpacted by human pollution and disturbance. That is, the regional reference site concept is meant to accommodate natural variations in biological communities due to bedrock, soils, and other natural physicochemical differences. Recognizing that pristine habitats are rare (even remote lakes and streams are subject to atmospheric deposition), resource managers must decide on an acceptable level of disturbance to represent an achievable or existing reference condition. Acceptable reference conditions will differ among geographic regions and States and will depend on the aquatic life use designations incorporated into State water quality standards.

Characterization of reference conditions depends heavily on classification of natural resources. The purpose of a classification is to explain the natural biological condition of a natural resource from the physical characteristics. Waterbodies vary widely in size and ecological characteristics, and a single reference condition that applies to all systems would be misleading. A classification system that organizes waterbodies into groups with similar ecological characteristics is required to develop meaningful reference conditions.

The best approach to classifying and characterizing regional reference conditions is determined by the estimated quality of potential reference sites that are available in the region. If a sufficient number of relatively undisturbed waterbodies exist (e.g., primarily forested watersheds), then it is possible to define watershed conditions that are acceptable for reference sites. If no reference sites exist, then reference conditions can be characterized based on an extrapolation of the biological attributes representative of the aquatic biota expected to be found in the region (see Gibson et al., 1996) or through other quantitative models (Hughes 1995). EPA sees the use of a regional reference condition as an important component and goal of State biological programs. The Agency also recognizes that other approaches, such as upstream/downstream sampling, may be necessary (U.S. EPA, 1990).

The Ohio Environmental Protection Agency has been very active in the development of biocriteria based on reference conditions. Ohio's experiences and methods may be useful to other States in developing their biological monitoring and biocriteria programs (see, for example, Ohio EPA, 1987, 1990). Florida DEP has developed a similar approach for defining reference conditions (Barbour et al., 1996); Arizona DEQ has oriented its reference condition by elevation (Spindler, 1996); and Maine DEC uses a statistically derived-decision model technique that is based on a knowledge of the ecology and expectations in the response to perturbation of the biological attributes to classify and assess its streams (Davis et al., 1993). For further information on the development and implementation of biological criteria and assessments, States should consult *Biological Criteria: National Program Guidance for Surface Waters* (U.S. EPA, 1990), *Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish* (Plafkin et al., 1989), and *Biological Criteria: Technical Guidance for Streams and Small Rivers* (Gibson et al., 1996).

C *A MULTIMETRIC APPROACH TO BIOASSESSMENT is recommended to strengthen data interpretation and reduce error in judgment based solely on population indices and measures.*

The accurate assessment of biological integrity requires a method that integrates biotic responses through an examination of patterns and processes from individual to ecosystem levels (Karr et al., 1986). The early conventional approach to using individual population measures has been to select some biological parameter that refers to a narrow range of changes or conditions and evaluate that parameter (e.g., species distributions, abundance trends, standing crop, or production estimates). Parameters are interpreted separately with a summary statement about the overall health. This approach is limited in that the key parameters emphasized may not be reflective of overall ecological health. The preferred approach is to define an array of metrics that individually provide information on each biological parameter and, when integrated, function as an overall indicator of biological condition. The strength of such a multimetric approach, when the component metrics are calibrated for a particular stream class, is its ability to integrate information from individual, population, assemblage, and zoogeographic levels into a single, ecologically-based index of water resource quality (Karr et al., 1986). The development of metrics for use in the biocriteria process can be partitioned into two phases (Barbour et al., 1995). First, an evaluation of candidate metrics is necessary to eliminate nonresponsive metrics and to address various technical issues (i.e., associated with methods, sampling habitat and frequency, etc.). Second, calibration of the metrics determines the discriminatory power of each metric and identifies thresholds for discriminating between "good" and"poor" sites. Known impaired sites are used to provide a test of discriminatory power. This process defines a suite of metrics that are optimal candidates for inclusion in bioassessments. Subsequently, a procedure for aggregating metrics to provide an integrative index is needed. For a metric to be useful, it must be (1) relevant to the biological community under study and to the specified program objectives; (2) sensitive to stressors; (3) able to provide a response that can be discriminated from natural variation; (4) environmentally benign to measure in the aquatic environment; and (5) cost-effective to sample. A number of metrics have been developed and subsequently tested in field surveys of benthic macroinvertebrate and fish assemblage (Barbour et al., 1995).

C *Assessment of HABITAT STRUCTURE as an element of the biosurvey is critical to assessment of biological response.*

Interpretation of biological data in the context of habitat quality provides a mechanism for discerning the effects of physical habitat structure on biota from those of chemical toxicants. If habitat is of poor or somewhat degraded condition, expected biological values are lowered; conversely, if habitat is in good condition (relative to regional expectations), high biological condition values are expected. Poor habitat structure will prevent the attainment of the expected biological condition, even as water quality problems are ameliorated.

If lowered biological values are indicated simultaneously with good habitat assessment rating scores, toxic or conventional contaminants in the system may have caused a suppression of community development. Additional chemical data may be needed to further define the probable causes (stressors). On the other hand, high biological metric scores in poor habitat could indicate a temporary response to organic enrichment, natural variation in colonization/mortality, change in predation pressures, change in food source/abundance, or other factors.

C *A standardized INDEX PERIOD is important for consistent and effective monitoring.*

The intent of a statewide bioassessment program is to evaluate overall biological conditions. The capacity of the aquatic community to reflect integrated environmental effects over time can be used as a foundation for developing bioassessment strategies (Plafkin et al., 1989). An index period is a time frame for sampling the condition of the community that is a cost-effective alternative to sampling on a year-round basis. Ideally, the optimal index period will correspond to recruitment cycles of the organisms (based on reproduction, emergence, and migration patterns). In some instances, an index period would be oriented to maximize impact of a particular pollutant source (e.g., high-temperature/low-flow period for point sources). Sampling during an index period can (1) minimize between-year variability due to natural events, (2) optimize accessibility of the target assemblages, and (3) maximize efficiency of sampling gear.

C *STANDARD OPERATING PROCEDURES and an effective QUALITY ASSURANCE PROGRAM are established to support the integrity of the data.*

The validity of the ecological study and resultant conclusions are dependent upon an effective QA Plan. An effective QA Plan at the onset of a study provides guidance to staff in several areas: objectives and milestones for achieving objectives throughout the study; lines of responsibility; accountability of staff for data quality objectives; and accountability for ensuring precision, accuracy, completeness of data collection activities, and documentation of sample custody procedures. Documented SOPs for developing study plans, maintenance and application of field sampling gear, performance of laboratory activities, and data analyses are integral quality control components of QA that can provide significant control of potential error sources.

- *A determination of PERFORMANCE CHARACTERISTICS of the bioassessment provides an understanding of the data quality for the assessment.*

  Perhaps the most important component in making bioassessments useful to water resource programs is the data quality of different assessment methods currently in use and the level of comparability among methods in performing an assessment. The comparability of methods should be judged by the degree of similarity in their performance characteristics (i.e., a performance-based approach) rather than by direct comparison of their respective scores or metric values (ITFM 1995, Diamond et al. 1996). To enable a sharing of data and results from various techniques that might be used by different agencies or other groups, some level of confidence in making an assessment must be established for each method based on the quality of data. This performance characteristic is precision, which is dependent upon the sampling methodology and the range in natural variation of the reference condition (note -- use of stream classification will increase precision).

  The ability to detect impairment also depends on the sensitivity of the method. In some cases, the desirable sensitivity level depends on how severe or subtle the impairment. For example, it does not require a very rigorous method to detect impairment following an extensive fish kill or algal bloom. It is the subtle impact areas that require some level of rigor that minimizes Type I and Type II errors in a judgment of condition.

  Based on preliminary information obtained from bioassessments conducted in Florida (Barbour et al. 1996a, Diamond et al. 1996), Ohio (Stribling et al. 1996), and New Hampshire (Stribling et al. 1994), quantitative criteria for precision and sensitivity can be set conservatively at “high” being less or equal to 20%, “moderate” being between 21 and 49%, and “low” being more or equal to 50%. High precision is equated to having low measurement error (coefficient of variation <20%) and sensitivity is the ability to detect small differences (<20% difference) between reference and the site being assessed.

- *AN IDENTIFICATION OF THE APPROPRIATE NUMBER OF SAMPLING SITES that are representative of a waterbody is an important consideration in evaluating biological condition.*

  The spatial array of sampling sites in any given watershed or region and the extrapolation of biological condition and water quality to areas beyond the exact sampling point must be established in any type of assessment. Two primary guidelines can be identified for extrapolating biological assessment data to whole watersheds. First, the structure of aquatic communities in lotic (flowing water) systems changes naturally with an increase in size of the stream. Thresholds in this continuum of change can be established through an analysis of regional databases. The biological condition at any particular site can only be used to represent upstream and downstream areas of the same physical dimensions and flow characteristics. Likewise, lake size will influence the number of sites needed to adequately characterize a lake or area of a lake. In small lakes, one site will generally be sufficient. In large lakes with multiple basins or in reservoirs with various zones (inflow, midsection, outflow), a site representative of each basin or zone may be needed.

A second consideration for site identification is the change in land use patterns along a stream gradient or lake shoreline. Changes from agricultural land use to urban centers, forested parkland, etc., would warrant different representative sampling sites. A waterbody with multiple dischargers may also require numerous sampling sites to characterize the overall biological condition of the waterbody.

Technical Support Literature

The Peer Review Team for ALUS recommended several technical papers to be used in support of specific technical issues associated with bioassessment. Information from these and other relevant literature will be incorporated into the revision of this chapter, pending comments and guidance from the Technical Experts Panel. The technical papers recommended by the ALUS Peer Review Team are as follows:

Cummins, K. W. 1988. Rapid bioassessment using functional analysis of running water invertebrates. In: T. P. Simon, L. L. Holst and L. J. Shepard (eds.). EPA -905-9-89-003. Proceedings of the First National Workshop on Biological Criteria. U.S. Environmental Protection Agency, Chicago.

Cummins, K. W. and M. A. Wilzbach. 1985. Field procedures for analysis of functional feeding groups of stream macroinvertebrates. Contribution 1611. Appalachian Environmental Research Laboratory, University of Maryland, Frostburg, Maryland.

Davis, W. S. and T. P. Simon (eds). 1995. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Hauer, F. R. and G. A. Lamberti (eds). 1996. Methods in Stream Ecology. Academic Press, San Diego.

Rosenberg, D. M. and V. H. Resh. 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman and Hall, New York.

3.3 Primary Contact Recreation Use

All States have recreational waterbodies with bathing areas, as well as less heavily used waterbodies with a designated use of swimming. In some States, nearly all waters are designated for swimming, although the great majority of waters are not used heavily for this purpose. States are asked to first target their assessments of primary contact recreation use to high-use swimming areas such as bathing beaches, a risk-based approach to targeting resources to protect human health.

3.3.1 Bathing Area Closure Data

States should acquire data on bathing area closures from State and local health departments and analyze them as follows.

A. Fully Supporting: No bathing area closures or restrictions in effect during reporting period.

B. Partially Supporting: On average, one bathing area closure per year of less than 1 week's duration.

C. Not Supporting: On average, one bathing area closure per year of greater than 1 week's duration, or more than one bathing area closure per year.

Some bathing areas are subject to administrative closures such as automatic closures after storm events of a certain intensity. Such closures should be reported along with other types of closures in the 305(b) report and used in making use support determinations if they are associated with violation of water quality standards.

3.3.2 Bacteria

States should base use support determinations on their own State criteria for bacteriological indicators.

EPA encourages States to adopt bacteriological indicator criteria for the protection of primary contact recreation uses consistent with those recommended in *Ambient Water Quality Criteria for Bacteria—1986* (EPA 440/5-84-002). This document recommends criteria for enterococci and *E. coli* bacteria (for both fresh and marine waters) consisting of:

- Criterion 1 = A geometric mean of the samples taken should not be exceeded, **and**

- Criterion 2 = Single sample maximum allowable density.

Many State criteria for the protection of the primary contact recreation use are based on fecal coliform bacteria as previously recommended by EPA (*Quality Criteria for Water—1976*). The previous criteria were:

- Criterion 1 = The geometric mean of the fecal coliform bacteria level should not exceed 200 per 100 mL based on at least five samples in a 30-day period, **and**

- Criterion 2 = Not more than 10 percent of the total samples taken during any 30-day period should have a density that exceeds 400 per 100 mL.

If State criteria are based on either of EPA's criteria recommendations outlined above (based on the 1976 or 1986 criteria), States should use the following approach in determining primary contact recreational use support:

A. Fully Supporting: Criterion 1 and Criterion 2 met.

B. Partially Supporting:

- For *E. coli* or enterococci: Geometric mean met; single-sample criterion exceeded during the recreational season, **or**

- For fecal coliform: Geometric mean met; more than 10 percent of samples exceed 400 per 100 mL.

C. Not Supporting: Geometric mean not met.

This guidance establishes a minimum baseline approach; should States have more restrictive criteria, these may be used in place of EPA's criteria. Please indicate when this is the case.

3.3.3 Other Parameters

In addition to pathogens, some States have criteria for other pollutants or stressors for Primary Contact Recreation. As noted by the ITFM, potentially hazardous chemicals in water and bottom sediment, ionic strength, turbidity, algae, aesthetics, and taste and odor can be important indicators for recreational use support determinations. The following guidelines apply where appropriate (i.e., where States have water quality standards for other parameters).

A. Fully Supporting: For any one pollutant or stressor, criteria exceeded in #10 percent of measurements.

B. Partially Supporting: For any one pollutant, criteria exceeded in 11 to 25 percent of measurements.

C. Not Supporting: For any one pollutant, criteria exceeded in >25 percent of measurements.

3.3.4 Special Considerations for Lakes

Trophic Status—

Trophic status is traditionally measured using data on total phosphorus, chlorophyll *a*, and Secchi transparency. As mentioned above, comparison of trophic conditions to natural, ecoregion-specific standards allows the best use of this measure.

In this context, user perception surveys can be a useful adjunct to trophic status measures in defining recreational use support. Smeltzer and Heiskary (1990) offer a basis for linking trophic status measures with user perception information. This can provide a basis for categorizing use support based on trophic status data. If user perception data are not collected in the State, extrapolations using data from another State, i.e., best professional judgment, might provide the opportunity to characterize recreational use support in a similar fashion.

<u>Pathogens</u>—

States should consider pathogen data in determining support of recreational uses. Guidelines above also apply to lakes.

<u>Additional Parameters</u>—

In addition to trophic status and pathogens, States should consider the following parameters in determining support of recreational uses:

- Frequency/extent of algal blooms, surface scums and mats, or periphyton growth
- Turbidity (reduction of water clarity due to suspended solids)
- Lake bottom siltation (reduction of water depth)
- Extent of nuisance macrophyte growth (noxious aquatic plants)
- Aesthetics.

3.4 Fish/Shellfish Consumption Use

Fish/Shellfish Consumption Advisory Data

A. Fully Supporting: No fish/shellfish restrictions or bans are in effect.

B. Partially Supporting: "Restricted consumption" of fish in effect (restricted consumption is defined as limits on the number of meals or size of meals consumed per unit time for one or more fish/shellfish species); or a fish or shellfish ban in effect for a subpopulation that could be at potentially greater risk, for one or more fish/shellfish species.

C. Not Supporting: "No consumption" of fish or shellfish ban in effect for general population for one more fish/shellfish species; or commercial fishing/shellfishing ban in effect.

In addition, the ITFM recommended specific indicators for assessing fish and shellfish consumption risks: levels of bioaccumulative chemicals in fish and shellfish tissue for fish and shellfish consumption, and, for shellfish only, paralytic shellfish poisoning (PSP)-type phytoplankton and microbial pathogens.

In areas where shellfish are collected for commercial or private purposes and removed to cleaner waters for depuration, the originating waterbodies should be considered Partially Supporting for Shellfish Consumption use.

3.5 Drinking Water Use

The following guidelines provide a framework for assessment of drinking water use support. These guidelines were developed by EPA in conjunction with the 305(b) Drinking Water Focus Group (DWFG), which consists of interested State and EPA personnel. EPA and States participating in the DWFG made it their goal to develop a workable set of guidelines that would serve to elevate the awareness of drinking water as a designated use within the 305(b) program, increase the percentage of waters assessed for drinking water use support, and enhance the accuracy and value of the assessments.

It was agreed by all parties involved in the development of these drinking water guidelines that no single template is suitable for every reporting State. The guidelines must incorporate flexibility and rely heavily on the judgment of the professional staff of each State’s public water supply supervision program to meet the challenges of assessing source waters for drinking water use support.

For purposes of the 1998 305(b) Water Quality Reports, States are asked to focus their assessments on water resources that support significant drinking water supplies. It is generally assumed that most States will initially focus their assessments on surface water resources; however, these guidelines are non-resource-specific and the framework may be applied to any waters within a State that are designated for drinking water use.

EPA and States participating in the DWFG discussed at length the issues and difficulties involved in assessing source waters for drinking water use support. EPA and these States recognize and fully accept that there will be significant variability in the information that States are able to provide in the 1998 305(b) reporting cycle. However, EPA expects that the direction of future reporting cycles will be evident, and that States will begin to develop plans and mechanisms to improve the overall accuracy and value of the assessments.

Key features of these guidelines include:

- assessment of State's water resources in phases over two 305(b) reporting cycles
- flexibility to perform assessments using a tiered approach
- identification of multiple data sources that may be used in the assessments
- assessment of water resources using a target list of contaminants reflecting the interests and goals of the State, and
- interpretation of data.

3.5.1 Prioritization and Phases of Source Water Assessment

EPA and the DWFG recognize that assessment of source waters for drinking water use support within the framework of the following guidelines is revised to achieve the key features listed above. EPA and the DWFG also recognize that assessment of the entire State's water resources for drinking water use support is a monumental task. To ease the burden, States may choose to perform drinking water use support assessments using a phased approach.

States may consider prioritizing their water resources and performing drinking water use support assessments for a limited percentage of their water resources. States are encouraged to expand their drinking water assessment efforts to include additional waters each subsequent reporting cycle. In this way, an increasingly greater percentage of waters will be assessed. Furthermore, this phased approach provides States with the opportunity to develop and implement plans and mechanisms for compilation, organization, and evaluation of drinking water data for improved reporting. EPA encourages States to set a goal of assessing drinking water use support for most of the State (approximately 75 percent of the waterbodies used for drinking water) by the year 2000.

For 1998, States are encouraged to set a priority for reporting results for waters of greatest drinking water demand. For these waters, States may elect to further prioritize with respect to vulnerability or other State-priority factors.

Identifying the presence of “treatment beyond conventional means” is one example of a technique that may be used to screen water resources for potential vulnerability and aid in prioritization of source waters for drinking water assessments. If “treatment beyond conventional means” is present (i.e., treatment beyond coagulation, sedimentation, disinfection, and conventional filtration), it may signify that the source water has been impacted to some degree and warrants more detailed investigation; however, it should be recognized that this information is generally not explicit, and therefore, neither the presence nor the absence of “treatment beyond conventional means” can be positively correlated to drinking water designated use support without additional investigation.

Prioritization of water resources for assessment may best be achieved in coordination with State professionals responsible for collecting and maintaining water quality data for sources of drinking water. It is generally these professionals that are most familiar with the data needed to assess drinking water designated use support and the conditions under which that data were collected. Their insight is integral to assuring the accuracy and value of these assessments.

3.5.2 <u>Tiered Approach for Source Water Assessments</u>

In addition to assessing only a limited percentage of State waters for drinking water use support, EPA and the DWFG encourage States to consider using a tiered approach in the assessments. A tiered approach accommodates the different types of data currently available to States with which to make an assessment and allows for differing levels of assessment.

Initially, States may use the most readily available information such as regional data, agency files, or other existing records or reports to conduct a preliminary assessment. As State programs develop and become more sophisticated, the preliminary assessments can be progressively upgraded through the incorporation of more detailed data (e.g., monitoring data). For 1998, EPA encourages States to provide a short narrative explaining how their assessments were performed and the level of detail incorporated into each assessment.

3.5.3 <u>Data Sources</u>

By instituting the tiered approach to conducting drinking water designated use assessments, EPA and the DWFG are acknowledging that data collection and organization varies among the States, and that a single data source for assessing drinking water designated use does not exist for purposes of the 1998 305(b) reports. EPA encourages States to use available data that they believe best reflect the quality of the resource. EPA is not asking States to conduct additional monitoring that does not fit in with other State priorities.

It is generally accepted that for purposes of the 1998 305(b) reports, States may need to be resourceful to acquire the data necessary to conduct preliminary assessments of source waters for drinking water designated use. States noted during the previous 1996 305(b) reporting cycle that the *Guidelines* placed heavy emphasis on the use of ambient water quality data. Frequently these data were not available and States defaulted to the use of finished water quality data. It was noted by many States that the default to finished water quality data might yield a jaded view of the source water quality.

EPA and the DWFG concur that the use of finished water quality data is not the best possible source of data for assessing source water quality; however, EPA and the DWFG also recognize the difficulties in obtaining data for use in drinking water assessments. By encouraging States to prioritize their water resources and perform drinking water use support assessments in a phased approach over two 305(b) cycles, EPA hopes that acquiring the necessary data will continue to become less difficult in time.

Within the numerous 1996 Amendments to the Safe Drinking Water Act (SDWA), the States are encouraged to use the Source Water Assessment Program (SWAP) to promote assessment of drinking water sources. EPA's August 1997 guidance suggests that States complete source water delineations and source inventory/susceptibility analyses for the public water supplies in the State within two years after EPA approval of the program. These assessments, when completed by the States, are an additional source of data for evaluating drinking water designated use and should contribute considerably to the assessment of drinking water quality.

For the 1998 305(b) reporting cycle, EPA is encouraging States to be resourceful in acquiring and using available data. EPA is not asking States to perform additional monitoring.

EPA and the DWFG identified several potential data sources that States might consider using in their 1998 assessments, including:

- Available ambient water quality data
- Untreated water quality data from public water supply (PWS) wells and/or surface water intakes[1]
- PWS drinking water use restrictions
- STORET database
- Independent water suppliers databases
- Source water assessments (SDWA 1996 Amendments)
- U.S. Geological Survey NAWQA studies
- Private water association studies
- Independent studies
- Other 305(b) use support impairments (e.g., aquatic life impairments).

States that have access to other data sources that can be used to assess source water quality for drinking water purposes are encouraged to use them if, in the judgment of the drinking water professionals, the data have undergone sufficient quality assurance/quality control checks.

Ideally, one or several of the above data sources will be available for States to use in assessing drinking water use support. However, lacking any of the above, States may have no choice but to default to the PWS compliance monitoring data required under the SDWA (i.e., finished water quality data). These data should only be used if the distinct source water can be identified (i.e., mixed systems do not qualify). Information on contamination-based drinking water use restrictions imposed on a source water may also be considered.

3.5.4 Contaminants Used in the Assessment

In many cases, the source of the data will determine the contaminants used in the assessment. For example, if a State has access to ambient

[1]States that designate for drinking water use only at the point of intake should assess an appropriate area of the source water for drinking water use support. This may require assigning an appropriate area around or distance upstream of the point of intake.

monitoring data, the assessment is limited to the monitored contaminants.

Each State should develop a target list of contaminants that best represents the State's assessment goals; this list may be based on monitoring or other sources of data. EPA and the DWFG recommend that States use the contaminants regulated under the SDWA as a starting point in developing their target list of contaminants (a list of the contaminants regulated under the SDWA and their associated maximum contaminant levels is provided in Appendix O). States are not expected to include all of the contaminants regulated under the SDWA as part of their target list.

EPA and the DWFG acknowledge that there are no specific guidelines or hierarchical structure to follow for developing a target list of contaminants for use in drinking water assessments and States must use their best professional judgment in the decision-making process. Important considerations include the availability and quality of data and the level of assessment States are prepared to make. To assist States in reducing the comprehensive list of contaminants regulated under the SDWA to a final, more manageable, grouping of contaminants, EPA and the DWFG recommend that States consider any of the following:

- MCL violations
- detections greater than the action trigger limits
- vulnerability studies
- occurrence data
- chemical waivers
- contamination-based drinking water use restrictions
- treatment beyond conventional means
- treatment objectives
- treatment processes
- treatment technique violations, and/or
- ambient turbidity levels.

EPA and the DWFG realize that the list of contaminants regulated under the SDWA is not an all-inclusive list and States may decide to add contaminants to their target group based on their best professional judgment. For example, States may choose to add contaminants that are not regulated under the SDWA but are of special interest or concern within the State (e.g., pesticides, herbicides, algae, phosphates).

3.5.5 <u>Data Interpretation</u>

EPA and the DWFG developed a framework to assist States in assigning use support categories based on data availability. As shown in Table 3-7, assessments can be based on actual monitoring data that are compared to water quality criteria (e.g., State-specific water quality standards or National Primary Drinking Water Regulations). If States do not have actual monitoring data available, finished water quality data and/or drinking water use restrictions could be used to infer source water quality. Use restrictions include:

- closures of source waters that are used for drinking water supply

- contamination-based drinking water supply advisories lasting more than 30 days per year

- PWSs requiring more than conventional treatment (i.e., other than coagulation, sedimentation, disinfection, and conventional filtration) due to known or suspected source water quality problems

- PWSs requiring increased monitoring due to confirmed detections of one or more contaminants (excluding cases with minimum detection limit issues).

3.5.6 Conclusion

Relatively few source waters have been adequately characterized for drinking water use support during the past 305(b) reporting cycles. EPA and States worked to develop a workable set of *Guidelines* that would serve to elevate the awareness of drinking water as a designated use within the 305(b) program, increase the percentage of waters assessed for drinking water use support, and enhance the accuracy and value of the assessments. These Guidelines provide a flexible framework for assessing drinking water designated use support. Using this framework is expected to result in better, more comprehensive assessments of source waters.

Table 3-7. Assessment Framework for Determining Degree of Drinking Water Use Support

Classification	Monitoring Data		Use Support Restrictions
Full Support	Contaminants do not exceed water quality criteria[a]	and/or	Drinking water use restrictions are not in effect.
Full Support but Threatened	Contaminants are detected but do not exceed water quality criteria[a]	and/or	Some drinking water use restrictions have occurred and/or the potential for adverse impacts to source water quality exists.
Partial Support	Contaminants exceed water quality criteria[a] intermittently	and/or	Drinking water use restrictions resulted in the need for more than conventional treatment with associated increases in cost.
Nonsupport	Contaminants exceed water quality criteria[a] consistently	and/or	Drinking water use restrictions resulted in closures.
Unassessed	Source water quality has not been assessed for contaminants used or potentially present.		

[a] For purposes of this assessment, EPA encourages States to use the maximum contaminant levels (MCLs) defined under the SDWA. However, if State-specific water quality standards exist, and constituent concentrations are at least as stringent as the MCL levels defined under the SDWA, State-specific water quality criteria can be used for assessment purposes.